Data Collection: Lessons of Cost-Benefit Analysis, Skepticism, and Legal Transparency

James X. Dempsey*

INTRODUCTION 127
I. ENDING COLLECTION PROGRAMS BASED ON A COST-BENEFIT ANALYSIS 127
II. WHEN PRIVACY IS AT STAKE, EFFICACY SHOULD ALWAYS BE AN ISSUE 129
III. A HEALTHY SKEPTICISM 130
IV. MAKING EFFICACY ANALYSIS ROUTINE 132
V. TRANSPARENCY 132
CONCLUSION 133

INTRODUCTION

In the realm of electronic surveillance, at least three lessons can be drawn from the experience of the last twenty years: (1) policymakers and the intelligence community can decide to *not* collect, or to stop collecting, information when its value does not outweigh the intrusiveness of its collection, even leaving aside questions of constitutionality or statutory authority; (2) policymakers, intelligence agency officials, and members of the public should be skeptical of claims—coming from both the intelligence community and the tech industry—about the value of big data analytics; and (3) intelligence collection programs can be described with some detail in legislation without compromising their effectiveness. For these lessons, I draw upon the experience of the now-defunct telephony metadata program under Section 215 of the PATRIOT Act and the much more successful collection activities under Section 702 of the Foreign Intelligence Surveillance Act.

I. ENDING COLLECTION PROGRAMS BASED ON A COST-BENEFIT ANALYSIS

One of the hard decisions in the intelligence field (a field characterized by hard decisions) is the choice to not collect certain information. At one level, intelligence is all about selection. It requires the constant funneling of oceans of information through finer and finer sieves to reach the point of knowledge or insight offered with some confidence. In the digital age, however, there is so much information available with such relative ease, with such powerful tools available to store and analyze it and such a faith in technology, that there may be a tendency to collect it "just in case." Once a collection program begins, it can be very

* Lecturer, UC Berkeley School of Law; Senior Policy Advisor, Program on Geopolitics, Technology and Governance, Stanford Cyber Policy Center; Member, Privacy and Civil Liberties Oversight Board (2012-2017).

difficult politically to end it, lest one be blamed for a subsequent attack that someone claims could have been prevented had the collection continued.

Nonetheless, policymakers and intelligence officials decided in the years after 9/11 to end the telephony metadata program because it posed risks to civil liberties without producing significant value. In other countries, the trend since 9/11—one that was accelerated by the Snowden leaks—has been to increasing bulk collection of data in the name of national security. The U.S., in contrast, may be the only country where the political branches publicly ended a bulk surveillance program. (In Europe, constraints on bulk collection have come from the regional human rights courts.)

The telephony metadata program involved the collection of call detail records—data indicating what number was calling what number, the date and time, and for how long any call lasted—for all telephone communications to, from, and within the U.S., including purely domestic calls. It was initiated in the weeks shortly after 9/11, based solely on an assertion of presidential authority. When that became untenable, the program was authorized by the Foreign Intelligence Surveillance Court, based on a reading of Section 215 of the PATRIOT Act. After the program was publicly disclosed by Edward Snowden, protracted debate resulted in its repudiation by Congress and the Obama Administration, and it was replaced by the USA FREEDOM Act of 2015 with a targeted program. Even that limited program was abandoned in 2019 when it proved impossible to implement, and the USA FREEDOM Act authority sunsetted.

The trajectory of the program, and the decision to end it, was based on a cost-benefit analysis. As a bulk collection program, the Section 215 program was at the high end of intrusiveness: by definition, it collected data on individuals suspected of no connection to terrorism and no wrongdoing of any kind. What ultimately doomed the program was its limited effectiveness, and therein lie several lessons.

Additionally, a second post-9/11 metadata program, involving internet communications, was ended by the intelligence community on its own initiative, without the forcing function of a leak. After its initial operation on the basis of Presidential fiat and then several years of operation under FISA court approval, the bulk collection of internet metadata was quietly terminated. Upon concluding that the program's value was limited, the NSA did not seek to renew it.

For me, the most salient legal problem with the telephony metadata program was that its operation bore almost no relationship to the text of the statute passed by Congress. Section 215 at the time authorized the FBI to obtain records, but in this case, the FBI obtained nothing: the records instead went to the NSA, which was not mentioned in the statute. The statute said any records obtained had to be handled pursuant to FBI guidelines; they were not, and instead were handled pursuant to NSA guidelines. The statute said that the records had to be relevant to "an authorized investigation," but they were not: the government argued instead that all the records obtained were relevant to all FBI investigations, current and future. Section 215 contemplated disclosure of data already in existence, but the

orders required prospective disclosure of records yet to be created. Prospective collection was covered by another section of FISA, which was not invoked for the program.

Beyond not fitting within the statute, the program raised Fourth Amendment concerns, falling into what remains the contested territory between *Smith v. Maryland*, 442 U.S. 735 (1979), and the subsequently-decided *Carpenter v. United States*, 138 S. Ct. 2206 (2018). Opponents of the program argued that such a comprehensive collection program could not be justified under the limited and targeted collection of third-party records allowed without a warrant under *Smith v. Maryland* and *United States v. Miller*, 425 U.S. 435 (1976). The government argued that the third-party doctrine holds that an individual has zero Fourth Amendment privacy interest in telephone records. An enormous volume of metadata, the government argued, still amounts to zero Fourth Amendment concern. So far, despite *Carpenter*, *Smith v. Maryland* still stands, but the privacy uneasiness with bulk collection remains.

The program also implicated the First Amendment right of association, a right recognized and protected by the Supreme Court in cases from *NAACP v. Alabama*, 357 U.S. 449 (1958), on. Although the NSA's telephone records program did not include an overt disclosure requirement of the type involved in such cases as *NAACP v. Alabama*, its operation similarly resulted in the compulsory disclosure to the government of information that could be used to identify individuals' lawful and protected political associations.

However, what really doomed the program was its lack of effectiveness.

II. When Privacy Is at Stake, Efficacy Should Always Be an Issue

Civil liberties advocates are wary, rightly so, of the argument that we must be willing to trade off some privacy in return for security. And certainly effectiveness should not rescue illegal conduct such as torture. But after 9/11, privacy advocates learned to deploy the trade-off analysis to the benefit of civil liberties. If a program intruding on privacy or other civil liberties is not effective, it should not be maintained regardless of whether it is constitutional or statutorily or judicially authorized.

When I first heard about the telephony metadata program, I assumed it would be effective. One of the most urgent tasks of the intelligence community after 9/11 was to find unknown individuals in the U.S. who might be planning further attacks. Those individuals were likely in communication with others, including others beyond our borders. To the extent that the intelligence community knew the identity and the communication identifiers of known or suspected terrorists here and abroad, surely it would be useful to track their communications with persons in the U.S. After all, weren't we bombarded all the time by claims of how big data could be used to identify relevant connections in all kinds of contexts?

To my surprise, and I suspect to the surprise of many others, the program was far less effective than the hype surrounding big data would have suggested. The Privacy and Civil Liberties Board, on which I served at the time, conducted the

most intensive review of any entity of the facts associated with the Section 215 program's value. The Board asked the intelligence community to provide information about every single instance in which Section 215 data proved useful. In the end, the PCLOB did not find a single instance involving a threat to the United States in which the telephone records program made a concrete difference in the outcome of a counterterrorism investigation. Moreover, the Board was made aware of no instance in which the program directly contributed to the discovery of a previously unknown terrorist plot or the disruption of a terrorist attack. The Board found that in only one instance over seven years had the program arguably contributed to the identification of an unknown terrorism suspect. In that case, the suspect was not involved in planning a terrorist attack and there was reason to believe that the FBI may have discovered him without the contribution of the NSA's program. Even in those instances where telephone records collected under Section 215 offered additional information about the contacts of a known terrorism suspect, in nearly all cases the benefits provided were minimal, generally limited to corroborating information that was obtained independently by the FBI.

Many inside the intelligence community came to the same conclusion. Some, indeed, seem to have long been skeptical of the program's value. Press reports indicated that the FBI, which was responsible for running to ground each lead developed by the program, was especially dubious as it turned up nothing on tip after tip.

In that context, the large-scale collection of telephony metadata was rightly abandoned, first replaced by a much narrower effort under the USA FREEDOM Act and then completely ended. NSA halted the program entirely in 2019 "after balancing the program's relative intelligence value, associated costs, and compliance and data-integrity concerns caused by the unique complexities of using these company-generated business records for intelligence purposes."[1]

III. A HEALTHY SKEPTICISM

The reality that the Section 215 program was not very effective contrasted sharply with both the initial claims of the government and the assumption of data omniscience that pervades our data-dependent economy and our tech-saturated personal lives.

Initial defense of the 215 program conflated its success with that of Section 702 collection. On June 18, 2013, in the first congressional hearing held after the Snowden leaks, the Director of the NSA testified that "[i]n recent years the information gathered from these programs provided the U.S. government with critical leads to help prevent over 50 potential terrorist events in more than 20 countries around the world. FAA 702 contributed in over 90 percent of these cases. At least 10 of these events included homeland-based threats. In the vast majority, business

1. Letter from Dan Coats, Dir. of Nat'l Intel., to Senators Richard Burr, Lindsey Graham, Mark Warner, and Dianne Feinstein 1 (Aug. 14, 2019), https://perma.cc/2BUU-9ZXG.

records, FISA reporting, contributed as well."[2] The Director later cited 54 cases in which the NSA bulk collection programs "contributed to our understanding, and in many cases helped enable the disruption of terrorist plots."[3]

By December 2013, the 50 or 54 cases had shrunk: Director Alexander spoke of only eight events in which Section 215 played a role in disrupting terrorist activity. By 2014, when the PCLOB completed its work, there was only one: the bulk metadata program had been useful in identifying a previously unknown terrorist suspect, Basaaly Moalin, who was not planning any attacks in the U.S., but was sending small amounts of money to Al-Shabaab, the extremist Somali militia with al-Qaeda ties.

The trajectory from 50 or 54 to eight to one may simply be a natural by-product of the processes that ensue when a government (or corporate) program generates a media storm, starting with an instinctive defensive reaction and then moving to more measured analysis. However, it should serve as a reminder of the need to be as skeptical of the government's claims in defending a program as one should be of the hyperbole of the program's critics.[4]

The lesson of skepticism has a corollary: policymakers and advocates concerned about civil liberties should not be afraid of inquiring into efficacy. Of course, a program may in fact be highly effective, and digging deeper into efficacy may end up strengthening the government's case for the program. However, in succumbing to the assumption of data's omniscience, civil liberties advocates may leave out of the public debate the most important reason for ending a program: it doesn't work.

A related lesson is that big data is hard to manage, at least if you care about adherence to rules intended to minimize unfair consequences for individuals. Prior to the USA FREEDOM Act, the Section 215 program was plagued by problems. One FISA judge wrote in 2009 that, from the inception of the program, "the NSA's data accessing technologies and practices were never adequately designed to comply with the governing minimization procedures."[5] The NSA's justification: among key personnel overseeing the program, there was never a complete understanding regarding what each individual meant by the terminology used in reports to the court. "Furthermore," said the NSA director, "from a technical standpoint, there was no single person who had a complete technical

2. *How Disclosed NSA Programs Protect Americans, and Why Disclosure Aids Our Adversaries: Hearing Before the H. Permanent Select Comm. On Intel.*, 113th Cong. (2013) (testimony of General Keith Alexander, Dir. Nat'l. Sec. Agency), https://perma.cc/5RKT-CQ6T.

3. General Keith Alexander, Dir. Nat'l. Sec. Agency, Remarks at the Armed Forces Communications & Electronics Association Conference (Jun. 28, 2013), https://perma.cc/24ZM-BT87.

4. A much harder question is why the program was not more successful. Susan Landau and Asaf Lubin have done impressive work answering that question without access to classified information. *See* Susan Landau and Asaf Lubin, *Examining the Anomalies, Explaining the Value: Should the USA FREEDOM Act's Metadata Program be Extended?*, 11 HARV. NAT'L SEC. J. 308 (2020).

5. DAVID MEDINE, RACHEL BRAND, ELISEBETH COLLINS COOK, JAMES DEMPSEY & PATRICIA WALD, REPORT ON THE TELEPHONE RECORDS PROGRAM CONDUCTED UNDER SECTION 215 OF THE USA PATRIOT ACT AND ON THE OPERATIONS OF THE FOREIGN INTELLIGENCE SURVEILLANCE COURT 52 (2014), https://perma.cc/FL9F-8HNQ [hereinafter PCLOB 215 REPORT]

understanding of the [program's] system architecture."[6] After the transition to the USA FREEDOM program, the data processing difficulties only became worse.

Constantly, we see a curious dynamic in discussions of data and data analytics in both commercial and government contexts. Proponents simultaneously claim that the use of data is fantastically revealing and simultaneously not threatening. Critics simultaneously argue that data is risky because it reveals so much and risky because it is prone to error. As in so many human efforts, these contradictory claims all have an element of truth. Data is simultaneously useful and hard to use; it is simultaneously revealing and misleading.

IV. MAKING EFFICACY ANALYSIS ROUTINE

As Susan Landau and Asaf Lubin have written, "understanding efficacy—the goals of a surveillance program and how well it achieves them—is essential to striking a balance between privacy and civil liberties on the one hand, and public safety and security on the other. Unlike the more expansive concerns over the balance between privacy and security, questions of efficacy are not philosophical or constitutional; they are rooted in pragmatism."[7]

There may still be too little consideration of effectiveness. In the PCLOB report on the Section 215 Program, Board member Elisebeth Collins Cook urged the intelligence community to spend more effort systematically examining the effectiveness of collection programs. Cook called on NSA and other members of the Intelligence Community to develop metrics for assessing the efficacy and value of intelligence programs, particularly in relation to other tools and programs, arguing that "[t]he natural tendency is to focus on the operation of a given program, without periodic reevaluations of its value or whether it could be implemented in more privacy-protective ways."[8] I am currently not in the position to know where things stand in this regard, but I would guess that the question of efficacy, and the application of cost-benefit analysis, remains undervalued.

Furthermore—relying again on the insights of Landau and Lubin—when both modes of communication and the nature of terrorism itself are changing, there is no reason to presume that any program, even if initially very effective, will remain efficacious 10 years after it was conceived.

V. TRANSPARENCY

The telephony metadata program was launched in secrecy. It was converted to FISA authorization in secrecy. During the program's lifetime, Section 215 was debated and reauthorized in congressional proceedings where some legislators knew what was happening and where most saw no connection between the words of the statute they were voting on and the conduct of the NSA and the telephone

6. *Id*. at 48.

7. Susan Landau and Asaf Lubin, *Examining the Anomalies, Explaining the Value: Should the USA FREEDOM Act's Metadata Program be Extended?*, 11 HARV. NAT'L SEC. J. 308, 351 (2020).

8. PCLOB 215 REPORT, *supra* note 5, at 217.

companies. When the program was leaked by Snowden, it immediately lacked credibility.

In contrast stands the FISA 702 program: in some ways, Section 702 collection is more intrusive, because it collects not just metadata but also the content of communications. On the other hand, the program is targeted, in that the government acquires only the communications associated with the specific accounts or devices of particular individuals, and those individuals can be targeted only if they are reasonably believed to be non-U.S. persons outside the U.S.

The PCLOB found that the information the 702 program collects had been valuable and effective in protecting the nation's security and producing useful foreign intelligence. Equally importantly, the board found that the text of the statute publicly outlined the basic structure of the program. The latter point is especially noteworthy given that the Section 702 program is extremely complex, involving multiple agencies, collecting multiple types of information, for multiple purposes.

The language of Section 702 is very complicated, involving certifications, authorizations, and directives (each of which means something different), three sets of procedures (targeting, minimization, and more recently, querying), multiple judicial reviews, and potential challenges and enforcement actions. However, PCLOB found that the program in practice matched the program described on the books. And the government described the program in considerable detail, clarifying that its application of the statute was actually at the narrower end of the range of possible interpretations.

The lesson is that intelligence programs can be described in some detail without compromising their effectiveness. Conversely, programs—especially domestic ones—conducted without clear statutory authority can survive too long without the accountability that every government program deserves.

CONCLUSION

A clear-eyed commitment to efficacy analysis by both proponents and opponents of any program. Skepticism about technology's reliability, again by both proponents and opponents. And transparency. These three lessons of intelligence that I draw from the post-9/11 experience seem especially urgent in the age of artificial intelligence that is now upon us.

***